The Astrophysical Journal, 796:100 (15pp), 2014 December 1 doi:10.1088/0004-637X/796/2/100
© 2014. The American Astronomical Society. All rights reserved. Printed in the U.S.A.

UPDATED *SPITZER* EMISSION SPECTROSCOPY OF BRIGHT TRANSITING HOT JUPITER HD 189733b

Kamen O. Todorov[1,2,5], Drake Deming[2], Adam Burrows[3], and Carl J. Grillmair[4]
[1] Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA; todorovk@phys.ethz.ch
[2] Department of Astronomy, University of Maryland at College Park, College Park, MD 20742, USA
[3] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[4] Spitzer Science Center, California Institute of Technology, Mail Stop 220-6, Pasadena, CA 91125, USA
Received 2014 April 3; accepted 2014 October 5; published 2014 November 12

ABSTRACT

We analyze all existing secondary eclipse time series spectroscopy of hot Jupiter HD 189733b acquired with the now defunct *Spitzer*/Infrared Spectrograph (IRS) instrument. We describe the novel approaches we develop to remove the systematic effects and extract accurate secondary eclipse depths as a function of wavelength in order to construct the emission spectrum of the exoplanet. We compare our results with a previous study by Grillmair et al. that did not examine all data sets available to us. We are able to confirm the detection of a water feature near 6 μm claimed by Grillmair et al. We compare the planetary emission spectrum to three model families—based on isothermal atmosphere, gray atmosphere, and two realizations of the complex radiative transfer model by Burrows et al., adopted in Grillmair et al.'s study. While we are able to reject the simple isothermal and gray models based on the data at the 97% level just from the IRS data, these rejections hinge on eclipses measured within a relatively narrow wavelength range, between 5.5 and 7 μm. This underscores the need for observational studies with broad wavelength coverage and high spectral resolution, in order to obtain robust information on exoplanet atmospheres.

Key words: eclipses – planetary systems – techniques: spectroscopic

Online-only material: color figures

1. INTRODUCTION

Over 1820 planets have been discovered so far in orbit around stars other than the Sun.[6] A sub-set of these exoplanets, often referred to as "hot Jupiters" owing to their sizes and orbital periods of less than 10 days, is particularly amenable to atmospheric studies via transit observations. Secondary eclipse broadband photometry of hot Jupiters has been very successful, and planetary emission has been measured via this method in numerous investigations (e.g., Charbonneau et al. 2005; Deming et al. 2005, 2011; Knutson et al. 2008; Stevenson et al. 2010; Lewis et al. 2013; Todorov et al. 2013). Secondary eclipse *spectroscopy* is much more difficult owing to photon limits—the narrowband secondary eclipse measurements have much lower signal-to-noise ratio than broadband photometry for a given target. The mid-infrared eclipse depth even for hot Jupiters is often below ~0.5%, making this technique only possible, with current technology, for transiting planetary systems such as HD 189733 and HD 209458, where the host star has high apparent brightness and the planet-star contrast is also relatively high, leading to measurements with relatively high signal-to-noise ratios. Secondary eclipse spectroscopy studies have so far been limited to observations in the mid-infrared with the *Spitzer Space Telescope* (Grillmair et al. 2007; Richardson et al. 2007; Grillmair et al. 2008; Swain et al. 2008a) and at shorter wavelengths with the *Hubble Space Telescope* (e.g., Swain et al. 2009a; Evans et al. 2013) and some premier ground-based facilities (e.g., Crossfield et al. 2012).

The secondary eclipse depth of a planet at a given wavelength is a measure of its emission; it is equivalent to the contrast between the planet and the star. Therefore, the eclipse depth as a function of wavelength can be used to study the planet's emission spectrum and therefore characterize its atmosphere. Investigations utilizing broadband secondary eclipse photometry have suggested that hot Jupiters have two classes of atmospheres—with and without a temperature inversion (e.g., Knutson et al. 2008; Machalek et al. 2009; Todorov et al. 2010). A planet with a "temperature inversion" is understood to have a layer in its upper atmosphere that is warmer than the layers below, while in non-inverted atmospheres the temperature decreases monotonically with altitude. Some examples of planets with evidence for inverted atmospheres include HD 209458b (Knutson et al. 2008), CoRoT-1b (Deming et al. 2011), XO-4b (Todorov et al. 2012), etc., while, e.g., HD 189733b (Charbonneau et al. 2008), TrES-1 (Charbonneau et al. 2005), and WASP-4b (Beerer et al. 2011) have evidence for a lack of temperature inversion in their atmospheres. In addition to these two classes, there are several hot Jupiters that have ambiguous mid-infrared photometry measurements, compatible with both inverted and non-inverted models. In some cases, e.g., CoRoT-2b (Deming et al. 2011) and HAT-P-12b (Todorov et al. 2013), the current models both with and without a temperature inversion fail to describe the observations completely, possibly owing to incorrect or incomplete assumptions about the chemical and physical properties of the modeled atmospheres.

The causes for the presence or absence of temperature inversions are currently not well understood. A possible explanation is that planets with inverted atmospheres have an additional chemical species in their upper layers that causes strong absorption at pressures below ~0.01 bar, which leads to extra heating of these atmospheric layers (Burrows et al. 2008; Fortney et al. 2008). The absorber has been proposed to be gas-phase TiO (Hubeny et al. 2003; Burrows et al. 2007, 2008; Fortney et al. 2006, 2008), but this is under debate (e.g., Zahnle et al. 2009; Knutson et al. 2010; Parmentier et al. 2013).

Emission spectroscopy of planets can also be used to study the climates of hot Jupiters. Since hot Jupiters have very short

[5] Current address: Institute for Astronomy, ETH Zürich, Wolfgang-Pauli-Strasse 27, 8093 Zürich, Switzerland.
[6] Jean Schneider, exoplanet.eu, as of 2014 September 1.

orbital periods, they are expected to synchronize their rotation period with their orbital periods within ∼1 Gyr, assuming zero orbital eccentricity (e.g., Jackson et al. 2008; Correia & Laskar 2010), and hence have permanent day- and nightsides. The transport of heat from the dayside to the nightside of the planet is directly affected by the strength and direction of the atmospheric currents. More efficient energy transfer leads to a cooler dayside. By measuring the dayside thermal emission via secondary eclipse observations, we can probe the heat transport efficiency. This parameter is, however, degenerate with the Bond albedo of the planet, which can also cool the dayside by reflecting the stellar flux. Constraints can be placed on the combination of both parameters by comparison with models (Cowan & Agol 2011).

The emission spectrum of a planet could also be used to detect and measure the abundances of various molecular species in the planetary atmosphere. However, most secondary eclipse measurements are based on broadband photometry, and the "spectra" constructed from such observations in multiple wavelengths have extremely low resolution. Despite this, there have been efforts to evaluate the abundances of molecules like CO_2, CO, CH_4, and H_2O based on broadband photometry (e.g., Stevenson et al. 2010). However, more recent retrieval efforts (e.g., Lee et al. 2012; Line et al. 2012, 2013; Barstow et al. 2013) have suggested that much higher spectral resolution and large wavelength coverage are often required to reliably measure molecular abundances, especially when considering other free parameters like the temperature-pressure structure of the atmosphere and the efficiency of heat transport to the nightside.

A previous study with the Infrared Spectrograph (IRS; Houck et al. 2004) on *Spitzer* by Grillmair et al. (2008) analyzed 10 time series emission spectroscopy data sets on HD 189733b with coverage between ∼5 and ∼14 μm and found a strong downturn in planetary emission below 10 μm, as well as a spectral feature associated with water vapor absorption.[7] In addition to these 10 secondary eclipse data sets, there are 8 in the same wavelength range and 4 between 21 μm and 40 μm that were also observed but have never been analyzed.

While there are dozens of planets observed via *Spitzer* photometry at the 3.6 and 4.5 μm bands during eclipse, only two planets have ever been observed via secondary eclipse time series spectroscopy with *Spitzer*—HD 189733b (Grillmair et al. 2007, 2008) and HD 209458b (Richardson et al. 2007; Swain et al. 2008a). This number will not increase, since after the cryogen on board ran out in 2009 *Spitzer* no longer has spectroscopic capabilities. The relatively high temperature of the planet and the relatively low temperature of the star result in deep eclipses in an IR-bright target, thus allowing for relatively high signal-to-noise measurements of the eclipse depths as a function of wavelength. (We list the planetary system parameters in Table 1.) In addition, in the past several years, our understanding for the systematic effects present in *Spitzer* time series observations has significantly increased. Motivated by the combination of these factors, we examine the available secondary eclipse spectroscopy data sets in a self-consistent and up-to-date manner in order to improve our understanding of the atmosphere of this hot Jupiter and test the Grillmair et al. (2008) results.

[7] Water detections have also been reported in the infrared transmission spectrum of the planet (Tinetti et al. 2007; Swain et al. 2008b; Birkby et al. 2013).

Table 1
Adopted Stellar and Planetary Parameters for HD 189733[a]

Parameter	Value
$M_\star$ ($M_\odot$)	$0.823^{+0.022}_{-0.029}$
$R_\star$ ($R_\odot$)	$0.766^{+0.007}_{-0.013}$
K_s (mag)[b]	5.541 ± 0.021
$T_{\rm eff}$ (K)	5090
$b_{\rm impact}$	$0.900^{+0.006}_{-0.010}$
$M_{\rm p}$ ($M_{\rm J}$)	$1.138^{+0.022}_{-0.025}$
$R_{\rm p}$ ($R_{\rm J}$)	$1.178^{+0.016}_{-0.023}$
P (days)	$2.21857312^{+0.00000036}_{-0.00000076}$
T_0 ($\rm BJD_{TDB}$)	$2453988.804144^{+0.000072}_{-0.000039}$
$a_{\rm p}$ (AU)	$0.03120^{+0.00027}_{-0.00037}$
e	$0.0041^{+0.0025}_{-0.0020}$

Notes.
[a] Values from Triaud et al. (2009), except for K_s and $T_{\rm eff}$ (Knutson et al. 2010).
[b] Two Micron All Sky Survey (2MASS) K_s magnitude of the star (from the Infrared Science Archive: http://irsa.ipac.caltech.edu).

In Section 2, we discuss the available data. Section 3 focuses on our analysis approach, while in Section 4 we discuss our results and compare them to theoretical models.

2. OBSERVATIONS

We analyze all 22 archival time series spectroscopy data sets on HD 189733b observed with *Spitzer*/IRS during secondary eclipse. These observations will remain unique until a new infrared telescope with spectroscopic capabilities in the mid-infrared becomes operational. The *Spitzer* Heritage Archive[8] (SHA) lists all data sets, including proprietary ones; thus, we are confident that our data include all IRS spectroscopy on HD 189733b during secondary eclipse. Our data cover a wavelength range between 5 and 40 μm at resolution $R \sim 100$. The information about the available data sets is summarized in Table 2. The principal investigator of all observations (program ID numbers 30473 and 40504) is Carl J. Grillmair.

3. DATA ANALYSIS

3.1. Spectroscopy Extraction

We start our analysis with the basic calibrated data (BCD) files produced by the *Spitzer* calibration software, version S18.18.0. Following the discussion in Eastman et al. (2010), we convert the UTC-standard (which includes leap seconds as often as every six months) timing information included in the headers to the continuous Barycentric Dynamic Time (TDB) standard. Each IRS image contains 128 × 128 pixels, and the spectra are dispersed in wavelength approximately in the direction of the pixel columns. The spatial information is along the pixel rows. For the 7.4–14 μm data, we elect to clip the first two and the last three pixel rows (rows 1–2 and 126–128, corresponding to the longest and shortest wavelengths, respectively), since they appear to be noisier and could be subject to unknown systematic effects. Similarly, we clip the first two and the last seven rows for the 5.0–7.5 μm observations, which cover only rows between 1 and 80 out of 128 on the images.

[8] http://sha.ipac.caltech.edu/applications/Spitzer/SHA/

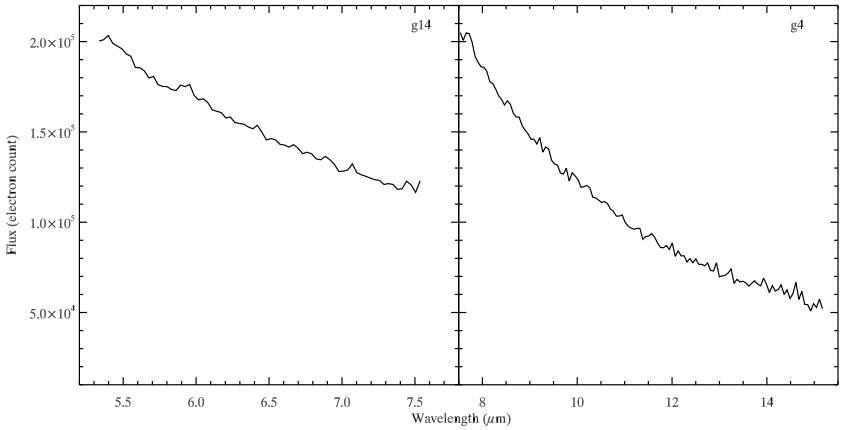

Figure 1. Example single spectra derived using optimal extraction (Horne 1986) for two data sets—g14 (left panel) and g4 (right panel). The next steps of the analysis follow the detected flux at a given wavelength as a function of time.

Table 2
HD 189733b Spectroscopic Observation Details

Data Set	AOR Key[a]	Observation Date	Wavelength Range (μm)	Exposure Time (s)	Spectra Count
g1	18245632	2006 Oct 21	7.4–14.0	14.7	900
g2	20645376	2006 Nov 21	7.4–14.0	14.7	950
g3	23437824	2008 May 24	7.4–14.0	61.0	280
g4	23438080	2008 May 26	7.4–14.0	61.0	280
g5	23438336	2008 Jun 2	7.4–14.0	61.0	280
g6	23438592	2008 May 31	7.4–14.0	61.0	280
g7	23438848	2007 Oct 31	7.4–14.0	61.0	280
g8	23439104	2007 Nov 2	7.4–14.0	61.0	280
g9	23439360	2007 Jun 26	7.4–14.0	61.0	280
g10	23439616	2007 Jun 22	7.4–14.0	61.0	280
g11	23440384	2008 Jun 9	5.0–7.5	61.0	280
g12	23440640	2008 Jun 4	5.0–7.5	61.0	280
g13	23440896	2007 Dec 7	5.0–7.5	61.0	280
g14	23441152	2007 Nov 6	5.0–7.5	61.0	280
g15	23441408	2007 Nov 11	5.0–7.5	61.0	280
g16	23441664	2007 Nov 9	5.0–7.5	61.0	280
g17	23441920	2007 Nov 24	5.0–7.5	61.0	280
g18	23442176	2007 Nov 15	5.0–7.5	61.0	280
g19	23439872	2007 Nov 4	13.9–21.3	122	140
g20	23440128	2007 Jun 17	13.9–21.3	122	140
g21	23442432	2007 Dec 10	19.9–39.9	122	140
g22	23442688	2007 Jun 20	19.9–39.9	122	140

Notes. [a] The Astronomical Observation Request (AOR) key that uniquely identifies the observation in the *Spitzer* Heritage Archive (http://sha.ipac.caltech.edu/applications/Spitzer/SHA/).

For each data set, we correct for energetic particle hits by following the value of a given pixel as a function of time. We replace pixel values that are at least 4σ away from a running median of width 5 with that median. In this manner, we correct about 0.8% (5.0–7.5 μm data), 0.4% (7.4–14 μm data), and 0.7% (14–40 μm) of the pixels in every image.

We employ optimal extraction (Horne 1986), implemented in IDL, to reduce the observed images to one-dimensional spectra. For each image, we estimate the background as a function of wavelength by fitting a third-order polynomial to every pixel row, i.e., along the spatial direction, excluding the region where the target was located. We determine the required polynomial order by experimenting with lower- and higher-order polynomials. Third order appears to be the lowest-order polynomial that consistently fits the background. While second-order polynomials often produce comparable fits, for some wavelengths they completely fail to match the background values in some region of the fitting domain. Thus, we elect third-order polynomials to fit the background. Again, for each pixel row, we subtract the corresponding background polynomial and then locate the peak of the source's brightness by fitting a one-dimensional Gaussian. Thus, we are able to follow any gradual curvature in the shape of the dispersed spectrum on the image. We make an initial estimate of the detected brightness by integrating the background-subtracted flux values of each row. The integration is centered on the estimated brightness peak for that row, and the integration range covers 3σ in each direction around the peak—therefore changing with wavelength. We use this initial spectral estimate as input for the algorithm by Horne (1986), which iteratively calculates the optimum one-dimensional spectrum extracted from each image. As per the standard practice, the algorithm only uses the 3σ span for the initial guess of the spectrum, but the iterations are performed over the whole spectral image. We show extracted sample spectra in Figure 1.

Initially, we integrate each derived spectrum over wavelength and construct "white light" curves. We present these for all data sets in Figure 2, except for the g1 light curve (taken on 2006 October 21), which we exclude, because only for these data was the telescope nodded between two positions approximately every 12 minutes (30 exposures). The nod was performed in order to facilitate background removal, but this exacerbated the systematic effects such as the ramp of intensity with time and the quasi-periodic apparent flux variation (see Section 3.2.1). As a result, despite the identical exposure times, the g1 data set white light time series exhibits higher noise than the g2 data set white light curve, where the telescope was not nodded. We also exclude the g19–g22 data sets from the analysis (covering wavelengths between 14 and 40 μm), since these observations have a very low signal-to-noise ratio and do not even allow for useful eclipse depth upper limit determination. Thus, we focus on the g2–g18 data sets.

3.2. Light Curve Analysis

3.2.1. Systematic Effects

We observe two major systematic effects that dominate the *Spitzer* IRS light curves. First, there is a positive correlation of observed brightness with time, often called the ramp (Deming et al. 2006). This effect is similar to the ramp reported by *Spitzer* photometric time series studies at 8 μm (e.g., Knutson et al. 2008; Agol et al. 2010; Todorov et al. 2010). It is thought

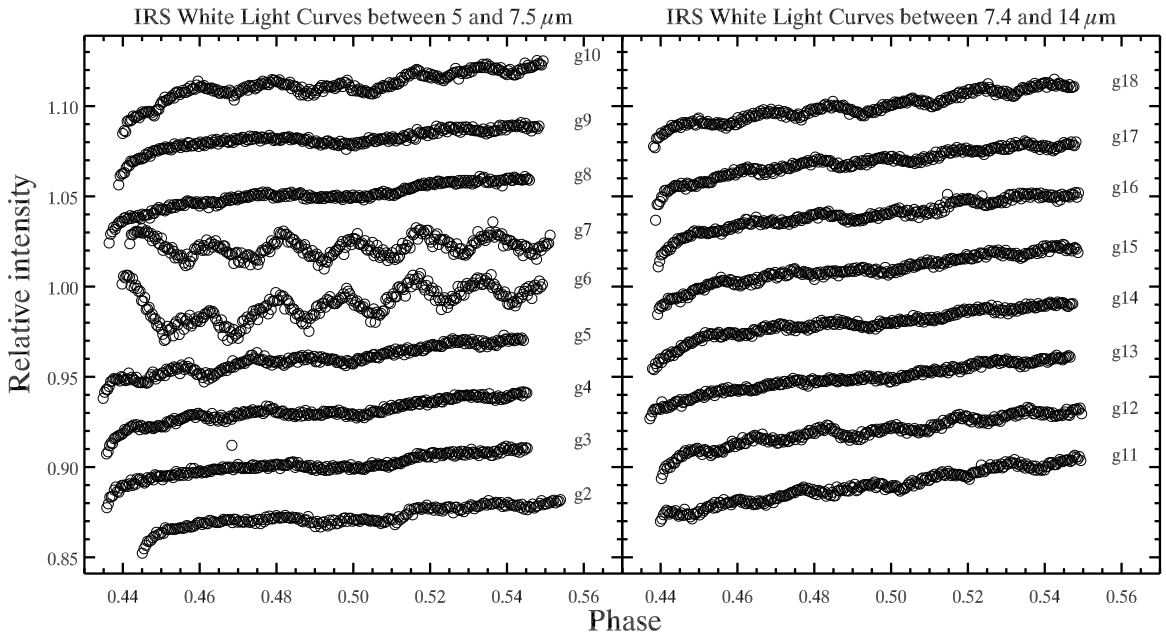

Figure 2. White light curves for all data sets that we include in our analysis, shown here, are calculated by integrating the extracted time series spectra from a given observation over wavelength. The resulting white light intensity as a function of time is normalized to one at the time of secondary eclipse and offset for clarity. We show both the 5–7.5 μm (left) and 7.4–14 μm (right) data. The abscissa here represents time in units of orbital phase. The timing information is extracted from the FITS file header as described in Section 3.1. We then convert the BJD_{TDB} times to phase by adopting the ephemeris from Table 1. The data sets are labeled based on their designations adopted in Table 2. One of the most obvious systematic effects is the apparent intensity oscillation with time. This is caused by a well-documented telescope pointing jitter (e.g., Grillmair et al. 2007; Deming et al. 2011). As the pointing shifts, the IRS slit is illuminated by various portions of the target PSF, resulting in changing apparent brightness. The systematic effects present in the data are discussed in Section 3.2.1.

that the cause for this ramp might be that photoelectrons are caught in quantum wells within the pixels and not read out at the start of the observation. As the observation progresses, most potential wells are filled and most photoelectrons are correctly read, causing the number of read-out electrons for a constant source to increase. Adopting this hypothesis, Agol et al. (2010) suggest a physically motivated toy model to account for the ramp,

$$\frac{F'}{F} = a_0 - a_1 e^{-t/\tau_1} - a_2 e^{-t/\tau_2}. \quad (1)$$

Here F' is the observed intensity before the correction for the ramp; F is the corrected intensity; a_0, a_1, a_2, τ_1, and τ_2 are free parameters; and t is time in arbitrary units. Here we use time in units of orbital phase, where $t = 0$ is the middle of primary transit and $t = 1$ occurs exactly one orbital period later. The coefficients in the first and second exponential terms are strongly correlated and degenerate. Thus, including all parameters in the Markov Chain Monte Carlo (MCMC) algorithm that we use to fit for the eclipse depths causes convergence problems. Therefore, we elect to neglect the second term in Equation (1),

$$\frac{F'}{F} = a_0 - a_1 e^{-t/\tau_1}. \quad (2)$$

This is justified since it results in only a marginal increase of the minimum χ^2 values achieved by the fits and a significant reduction of their Bayesian information criterion value (BIC; Schwarz 1978).

We compare this to several other ramp correction functions that are not physically motivated but have been used in past studies or are simple extensions of the single exponential function. We explore the log-linear (e.g., Todorov et al. 2010), the log-quadratic (e.g., Charbonneau et al. 2008), the third-order polynomial (Grillmair et al. 2007), which was used in the older analysis on part of our data, and the single exponent with a linear or quadratic functions added:

$$\frac{F'}{F} = a_0 - a_1 t - a_2 \log(t), \quad (3)$$

$$\frac{F'}{F} = a_0 - a_1 t - a_2 t^2 - a_3 \log(t), \quad (4)$$

$$\frac{F'}{F} = a_0 - a_1 t - a_2 t^2 - a_3 t^3, \quad (5)$$

$$\frac{F'}{F} = a_0 - a_1 e^{-t/\tau_1} - a_2 t, \quad (6)$$

$$\frac{F'}{F} = a_0 - a_1 e^{-t/\tau_1} - a_2 t - a_3 t^2. \quad (7)$$

The symbols for these expressions are defined as for Equation (1). All of the functions tested produce similar eclipse spectrum shapes, but the single exponential function typically produces the lowest BIC values in the individual wavelengths, especially at wavelengths longer than $\sim$8 μm. Thus, we elect to adopt it for the correction of the ramp in our analysis. While the choice of ramp correction function does not impact the overall shape of the planetary emission spectrum, it affects the absolute scale of the eclipse depths typically by $\sim$20%. We discuss this effect further in Section 4.1.

The second systematic effect is a quasi-periodic variation in detected flux with a period of about 60 minutes. This is caused by the well-known (e.g., Grillmair et al. 2007; Knutson et al. 2008; Deming et al. 2011) *Spitzer* pointing oscillation. This has been traced to the periodic working cycle of a heater that keeps a battery within its operating temperature.[9] In photometric data,

[9] http://ssc.spitzer.caltech.edu/warmmission/news/21oct2010memo.pdf

the change in observed brightness of the target is caused by the variable sensitivity across the surface of a single pixel. For IRS *spectroscopy* data, however, the slit drifts with the pointing oscillation and samples parts of the stellar point-spread function (PSF) that have different brightness. To put this in perspective, while the IRS slit is 3.″6 wide, the angular resolution of the telescope varies between 1.″5 and 4.″2 in the 5–14 μm range.[10] This effect is easily seen in the white light time series shown in Figure 2. We refer to it as the "sawtooth" effect, following the shape of the decorrelation function adopted by (Grillmair et al. 2008)—an asymmetric triangular sawtooth function with constant amplitude with time. However, in Figure 2 it is evident that the amplitude of the sawtooth correction may be variable throughout a single secondary eclipse observation. Thus, we explore alternative sawtooth removal methods.

Ideally, if precise pointing information was known for the telescope during the time of the observations, the pointing oscillation, and thus the shape of the sawtooth correction function, could be reconstructed from first principles. However, the pointing information recorded in the FITS file headers shows no pointing jitter, nor does the raw spacecraft telemetry that was graciously provided to us by the Spitzer Science Center HelpDesk and Sean Carey. It is possible that the temperature changes due to the operation cycle of the heater responsible for this effect cause the observatory star trackers to begin to lose and then regain their alignment with the telescope's boresight. Thus, despite the real pointing oscillation, the star trackers record no shift in the direction at which the main mirror is pointed. For the same reasons, the spacecraft gyroscopes also do not provide any useful information.

Another option is to centroid on a star whose image falls on one of the IRS peak-up arrays and is imaged during every IRS exposure. Unfortunately, there are no fortuitous high signal-to-noise ratio observations of point sources recorded in the IRS peak-up images.

However, there are multiple spectroscopic light-curve observations of HD 189733 for each wavelength range. Since the amplitude and phase of the sawtooth variation are unrelated to the astronomical observations, they are different for the different light curves. Hence, by stacking the white light curves together, we are able to nearly average out the sawtooth variations. When combining, we omit the g6, g7, and g10 data sets as a result of their relatively high sawtooth amplitudes. This results in high signal-to-noise ratio light curves, in which the dominant systematic effect is the ubiquitous *Spitzer* "ramp" and the sawtooth contribution is minimized. We show these curves in Figure 3.

Measuring the broadband eclipses and the ramps in the white light *stacked* light curves allows us to subtract them from the *individual* white light curve observations (shown in Figure 2), which leaves us with 17 residual curves, corresponding to each eclipse observation. We smooth these to extract the shape of the sawtooth correction that we need for each data set. The pointing oscillation of the telescope does not depend on wavelength, but the size of the stellar PSF compared to the size of the slit does. Therefore, the shape of the sawtooth correction is independent of wavelength, but the amplitude of the sawtooth *does* vary with wavelength. Thus, if we multiply the derived sawtooth correction function by a wavelength-dependent scaling factor, we can apply it to individual single-wavelength light curves. An advantage of this method is that it makes no assumptions about

[10] http://irsa.ipac.caltech.edu/data/SPITZER/docs/irs/irsinstrumenthandbook/4/.

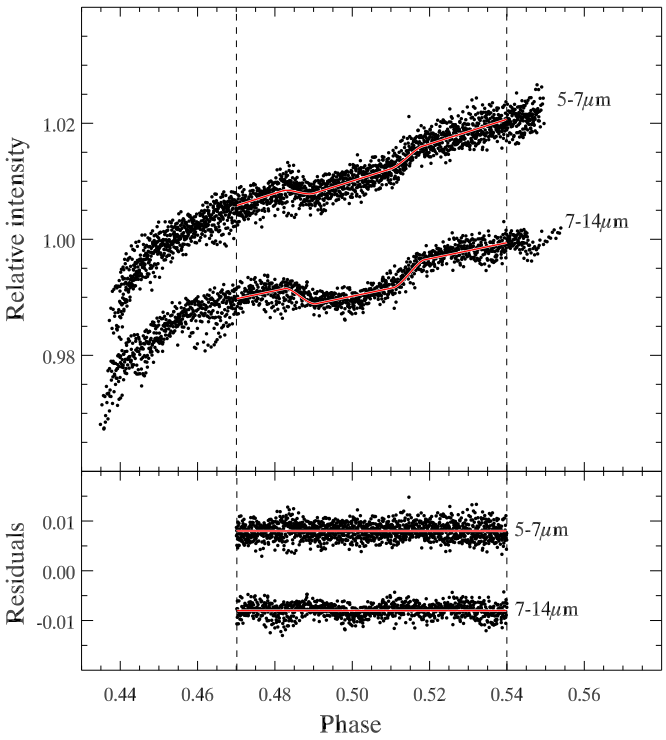

Figure 3. Upper panel: stacked white light curves in the 5–7 and 7–14 μm ranges, normalized to unity and arbitrarily offset for clarity. In the combined white light curves, the sawtooth effect is largely canceled. The red lines indicate the best-fit results from our MCMC analysis, including a ramp with time and a secondary eclipse, but not the sawtooth. The phase coverage of the red line and the dashed lines denote the data points that were included in the MCMC fits. Lower panel: residuals achieved by subtracting the best-fit eclipse and ramp models from the stacked white light curves, again arbitrarily offset for clarity. The red lines indicate the zero levels for the corresponding white light residuals.

(A color version of this figure is available in the online journal.)

the shape of the sawtooth. However, we lose any information about the variability of the planetary emission, since we stack the white light curves, implicitly assuming that the broadband eclipse depths are equal in all light curves. Since Agol et al. (2010) place a 1σ upper limit on HD 189733b's variability at 8 μm of 2.7%, any variability is likely to be below the *Spitzer*/IRS detection limit. Thus, the constant flux assumption is justified for this planet.

The lower panel of Figure 3 still shows some residual red noise, especially in the 7–14 μm range. This could bias the fit toward slightly deeper or shallower eclipse measurements than the "true" value. Since we are subtracting the best-fit eclipses from the stacked white light curves from the individual light curves, this implies including a small constant positive or negative eclipse in the sawtooth correction, leading to a small offset in the measured secondary eclipses (i.e., planetary intensity) in individual wavelengths. We stress that this effect, if it occurs, should have no impact on the shape of the planetary spectrum. We describe the details of our sawtooth correction algorithm in detail in the next section.

3.2.2. Fitting Procedure

In order to stack the light curves correctly, we convert the observation time for each spectrophotometric data point in all data sets to units of orbital phase using the ephemeris given in Table 1 (Triaud et al. 2009). We then normalize each light curve so that its average flux is unity during the expected time of secondary eclipse. Next, we simply combine all white light

photometric measurements in a single data set, where the data points are ordered by their calculated orbital phase. Here we bin the points from the g2 data by four points, because the exposure times for this observation were about four times shorter than for the other data sets (see Table 2). For the stacked white light curves, we adopt the following expression as a model to fit to the data:

$$I(t) = a_0 - a_1 e^{-t/\tau_1} + d_1 M_c. \quad (8)$$

Here d_1 is the eclipse depth and a_0, a_1, and τ_1 are the ramp parameters from Equation (2). The eclipse shape, M_c, centered on a given central phase, c, is based on the Mandel & Agol (2002) model but without limb darkening (since the planet is behind the star, the bottom of a secondary eclipse light curve is flat). The duration of the eclipse and ingress/egress is fixed, based on the planetary and stellar radii, eccentricity, and orbital period adopted in Table 1. We set a_0, a_1, τ_1, d_1, and c to be the free parameters for our stacked white light curve fits.

Not all data sets were observed starting at the exact same orbital phase, and hence the steep part of the ramp occurs at slightly different phase for each light curve. This makes stacked light curves noisy at phases earlier than ~0.47. In addition, not all data sets cover orbital phases higher than ~0.54, causing the stacked white light curves to be vulnerable to poor sawtooth cancellation for phases above this threshold. Thus, we only use data with phases between 0.47 and 0.54 for the fit of the stacked white light curves and the stacked light curves at individual wavelengths described in forthcoming sections.

We implement an MCMC fitting routine following the algorithm outlined by Ford (2005, 2006). We perturb only one, randomly selected, free parameter at a time. We perform 6×10^6 MCMC steps ($\sim 10^6$ steps per parameter). Of the total length of the chain, we drop the initial 10^6 steps as "burn-in" time, required for the chain to converge. Before running the long chain, we run several shorter chains in order to optimize the widths of Gaussian distributions that are used to determine the size of the parameter perturbation. We elect these widths to be such that the acceptance rate of the new parameter value is between 35% and 55% in order to optimize the efficiency of convergence (Ford 2006). The histograms of the parameter runs closely resemble Gaussians; therefore, we adopt the mean values for the eclipse depth and the eclipse central phase as the best-fit values for these MCMC fits. The MCMC states that result in the minimum χ^2 values have eclipse depths typically within 3%, or much less, of the mean histogram value. Therefore, the choice of "best value" does not change our final results. The best-fit eclipse depth in the 5–7 μm range data is 0.216% with a central phase of 0.50014. For the 7–14 μm range stacked white light curve, the eclipse depth is 0.370% and the central phase is 0.50060. Just due to the light-travel time delay of the HD 189733b system, and adopting $e \cos \omega = 0$ (Agol et al. 2010), the expected central phase is 0.50016. In addition, Agol et al. (2010) detect an additional 38 ± 11 s (0.00020 ± 0.00005 in units of phase) delay that they attribute to the hottest point of the dayside of the planet lagging behind along the orbit compared to the substellar point. Hence, our white light central phase measurements are compatible with previous studies, after allowing for the fact that at different wavelengths the delay due to hot spot offset may be different, since we are probing different atmospheric altitudes.

In order to estimate the sawtooth correction function for each individual data set, we use the same MCMC algorithm as above to fit a ramp (but still no sawtooth) to the individual white light curves shown in Figure 2. This time, we hold the eclipse depth and central phase fixed to the values derived from the corresponding stacked white light curves. As above, we use the ramp parameters from Equation (2). We subtract the best-fit "ramp-and-eclipse" model from the white light curve, and we are left with residuals that represent the sawtooth function combined with the photon noise. In order to eliminate the photon noise and estimate the sawtooth, we bin the residuals by six. We choose this factor empirically—bigger bins smooth the sawtooth curve too much and degrade the fits, while smaller bins are dominated by the photon noise. We then utilize the cubic spline interpolation IDL routine *spline* to evaluate the sawtooth function at the orbital phases when the data were actually observed. A sample sawtooth function is presented in Figure 4.

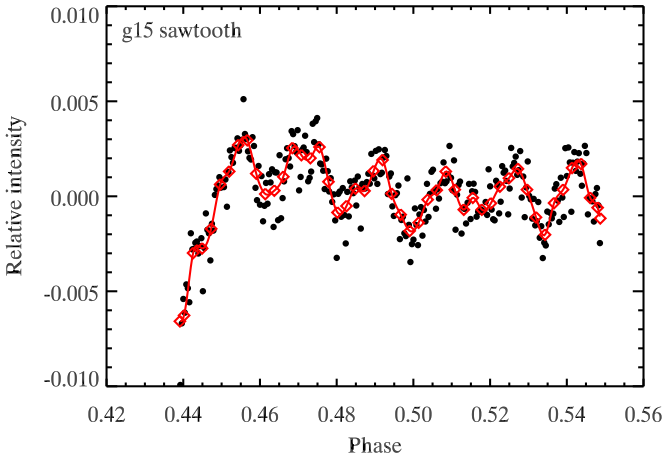

Figure 4. Residuals after subtracting the ramp and eclipse from the g15 white light curve (black points). Since this data set covers the 5–7 μm range, we fixed the eclipse depth at 0.216% and the central phase at 0.50014. The residuals were binned by six (red diamonds), and a cubic spline interpolation was used to estimate the final sawtooth function at the observed orbital phases (solid red line).

(A color version of this figure is available in the online journal.)

The light curves derived from a single pixel row (the wavelength dispersion direction is along the image columns) have a very low signal-to-noise ratio. To improve this, we bin the light curves in bins of width three pixels in *wavelength*, such that a given wavelength channel is the combination of three image rows, not just one. In this way, we are left with 24 channels between 5 μm and 7 μm, and 42 channels between 7 μm and 14 μm, per data set. We stack all light curves in a given wavelength channel in the same way that we stack the white light curves (but this time including the g6, g7, and g10 data sets). Again, the g2 data points are binned by a factor of four to account for their shorter exposure times. Since we have estimated the value of the sawtooth function for each data set and for each orbital phase, we stack the sawtooths in the same way that we stack the light curves—by simply combining the data in a single data set and ordering the points in it by their orbital phase. In this way each photometric point in the stacked light curves will be corrected by the corresponding sawtooth value of its original data set.

The amplitude of the sawtooth correction is dependent on wavelength, as discussed in Section 3.2.1. To account for this, we introduce a new free parameter, the sawtooth scale, ξ. We assume that ξ depends only on wavelength and that there is a single value of ξ that applies to all data sets at a given wavelength channel, i.e., that it is independent of time. Because the non-stacked light curves for any given channel are still relatively noisy, this assumption is difficult to test in practice. However, it

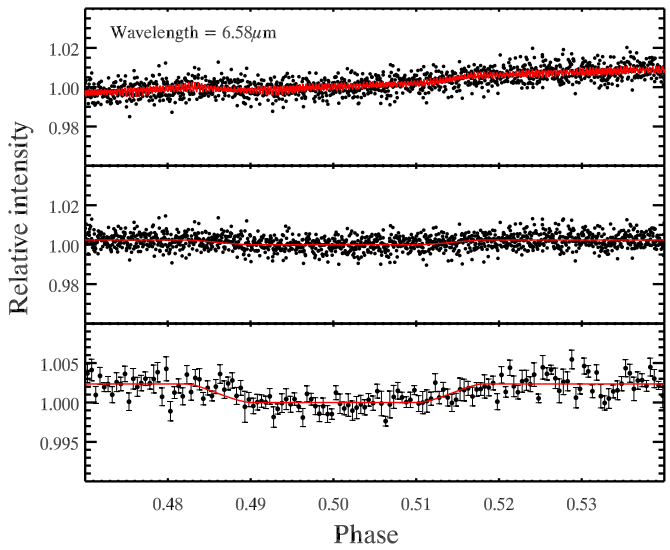

Figure 5. Raw (top panel, black points) stacked light curves at 6.58 μm based on all nine available data sets that contain this wavelength channel are fitted with a model (red line) including the eclipse depth, the *Spitzer* ramp, and the sawtooth correction. In the middle panel, the light curve (black points) has been corrected for the systematic effects, with the best-fit eclipse model shown with the red line. The bottom panel is similar, but here the corrected light curve is binned, each bin with length of 0.0005 in units of phase (1.6 minutes, or ~10 data points). The vertical axis is rescaled to emphasize the eclipse, which is clearly visible in the binned light curve. We exclude the intensity measurements with orbital phases below 0.47 and above 0.54, as discussed in Section 3.2.2.

(A color version of this figure is available in the online journal.)

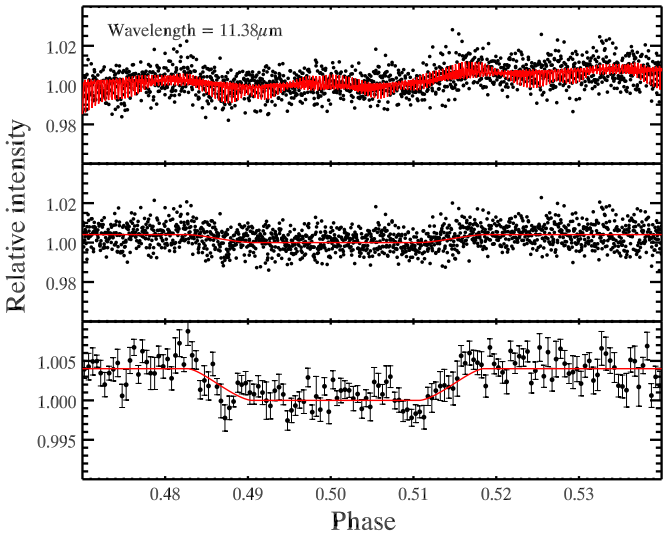

Figure 6. Similar to Figure 5, but here we show the 11.38 μm light curve, composed using all eight data sets that include this wavelength. The bins in the bottom panel, again, cover 0.0005 in units of phase, or 1.6 minutes, but for these data this corresponds to ~11 points per bin.

(A color version of this figure is available in the online journal.)

is reasonable, since the width of the target PSF compared to the slit width is what determines the sawtooth amplitude resulting from a given pointing shift, and this is independent of time and only depends on the wavelength. Variable PSF width can change this, but *Spitzer* is on an Earth-trailing orbit and is far more thermally stable than, e.g., the *Hubble Space Telescope* on its low-Earth orbit. Therefore, it is not surprising that we see no significant changes in the point response function of the telescope that might indicate a change in the relative sizes of the slit and the PSF at a given wavelength.

3.2.3. Best-fit Depths

We perform both MCMC and Prayer Bead Monte Carlo (PBMC; Gillon et al. 2007) fits for the stacked light curves at each of the 56 wavelength channels. Our MCMC fits utilize the same algorithm described in Section 3.2.2, except here we include the sawtooth correction. We model the light curves with the following expression:

$$\frac{F'}{F} = a_0 - a_1 e^{-t/\tau_1} + \xi S_s + d_1 M_c. \quad (9)$$

S_s represents the stacked sawtooth correction function. The free parameters in this fit are a_0, a_1, τ_1, ξ, and d_1. We fix the central phase to the best-fit value from the stacked white light curve fits since it is approximately independent of wavelength.[11] For each wavelength channel, we record the MCMC free parameter runs and create a histogram of the values for each parameter that closely resemble Gaussian distributions. We adopt the mean of the histogram to be the "best fit." As before, using the minimum χ^2 eclipse depths does not change our "best-fit" values by more than ~3%, much less than our uncertainties, and the choice of best-fit value has no impact on the final results. We show typical raw and corrected light curves for two wavelength channels, along with the best-fit MCMC models for the eclipse and the systematic noise in Figures 5 and 6.

It is possible that our sawtooth correction method does not remove the sawtooth perfectly, and there could be additional, less important, systematic effects that we have not taken into account. Therefore, there may be residual red noise in the systematics-corrected light curves (e.g., middle and lower panels of Figures 5 and 6). Similarly to the photometric analyses in, e.g., Désert et al. (2011), Deming et al. (2011), and Todorov et al. (2012, 2013), in order to account for the possible influence of the residual correlated noise, we perform PBMC fits for the light curves observed at each wavelength channel. The PBMC is based on simulating additional data sets by subtracting the best-fit residuals from the observed data. The residuals are typically shifted over by a fixed number and added back to the best fit. This creates a simulated observation where the red noise from the original data is preserved. Simulating multiple data sets in this manner and fitting the eclipse depth model to them allows us to quantify the impact of correlated noise on our results, unlike the MCMC fitting routine, which assumes Gaussian uncertainties.

For the PBMC fits, we adopt the MCMC model realization with the smallest χ^2-statistic to be the best-fit model. Shifting the residuals of the stacked light curves with respect to the best-fit model is inappropriate in this case, because each observed eclipse within the stacked light curve may have its own associated red noise, independent of that of the other observations. Therefore, instead of shifting the stacked light curve residuals by one at every PBMC iteration, we calculate the residuals of each individual light curve at a given wavelength separately and shift each of them by a random number. Then we add the shifted residuals back to the best-fit model and stack the thus-simulated light curves the same way as the original data. Performing an MCMC fit to the simulated stacked light curves

[11] The hot spot offset from the substellar point inferred by Knutson et al. (2007) and Agol et al. (2010) at 8 μm, to which the eclipse central phase delay is attributed, may be different at different atmospheric altitudes. This, combined with the fact that different observing wavelengths probe different layers of the atmosphere, means that the apparent delay of the secondary eclipse central phase may be dependent on wavelength. However, this effect is most likely very subtle and can safely be ignored here.

is extremely computationally expensive. Therefore, we fit the model using the IDL *mpfit* package for nonlinear least-squares fitting (Markwardt 2009). The free parameters used here are the same as the ones used for the observed data MCMC fit. In this way, we simulate 10,000 stacked light curves per wavelength channel, and this allows us to evaluate the cumulative effect that the residual correlated noise may have on the eclipse depth results. We create histograms from the eclipse depths from the PBMC runs for each wavelength. These are typically close to Gaussian in shape, but often with some asymmetry as expected for data where small amounts of red noise remain uncorrected. Since the astrophysical white noise is unrelated to the telescope systematic noise, any of the simulated PBMC realizations could have been the *observed* data. Therefore, we choose to adopt the medians of these histograms instead of the MCMC best-fit values as the "best-fit" points in the histograms.

3.2.4. Uncertainty Estimates

We fit each MCMC histogram with a Gaussian function and adopt its standard deviations to be the uncertainty of the parameter at the given channel. We define the 1σ PBMC uncertainties as the region that covers 68% of the histogram centered on the median simulated-data eclipse depth. We select this definition because it allows for efficient comparison on an even footing with the results from the MCMC Gaussian uncertainties. We compare the uncertainties from the PBMC and MCMC estimates of the eclipse depths in all wavelength channels. We find that the PBMC uncertainties are, in all cases, larger, but the PBMC "best values" (the histogram medians) are always well within 1σ of the MCMC best fits. The MCMC uncertainties are smaller than the PBMC ones by between 14 and 69%, but typically by about 50%.

Both the MCMC and PBMC errors give good estimates of how the magnitude of the error bars varies with wavelength, due to changing photon counts per exposure per wavelength – their wavelength dependence is very similar. However, they are imprecise estimates of the errors in an absolute sense—the MCMC uncertainties do not take residual red noise into account, while the PBMC uncertainties give large uncertainties to account for the uncertainty in the absolute eclipse depths, even though they do not account for systematic effects, like the choice of ramp correction function (Sections 3.2.1 and 4.1). Thus, the PBMC uncertainties overestimate the uncertainty in the difference between two wavelengths. Since the overall absolute contrast may vary depending on the choice of a ramp function, we focus on the point-to-point changes in the spectrum. Therefore, we use the PBMC and MCMC uncertainties to describe the wavelength variation of the errors, but we exploit the nature of the spectrum to scale all of the errors accurately to reflect their correct wavelength-to-wavelength uncertainty.

For any measurement, the best way to estimate the real magnitude of the errors is to repeat the measurement many times independently and compare the different results. This is impractical to do for eclipse photometry (except in a few cases, e.g., Agol et al. 2010), but the nature of low- to moderate-resolution spectra can be exploited as a proxy for independent measurements. Because the spectrum changes only slowly with wavelength compared to the precision of our data, the eclipse depth should be closely the same at adjacent wavelengths. Hence, we use the wavelength-to-wavelength differences in derived eclipse depth to scale the errors in an absolute sense.

To scale the uncertainties, we adopt the following procedure: first, we take the HD 189733 star-planet contrast model from Burrows that we present in Section 4.5 and, and we bin it to the resolution of our extracted IRS spectrum. We find that the binned model varies smoothly from spectroscopic channel to channel. We express the wavelength dependence of the uncertainties (longer wavelengths yield higher uncertainties) by taking the average of the PBMC and the MCMC uncertainties for a given wavelength, ζ_i. We add random Gaussian noise to the binned Burrows model, with $1\sigma = k\zeta_i$ at the respective wavelengths, where k is a scaling factor. In this way, we simulate observations for a range of values for k, varying it in steps of 0.01. We estimate the point-to-point scatter of the observed and simulated eclipse spectra by taking the standard deviations of their derivatives. We find that they are closest for $k = 0.887$. Thus, as final uncertainties, we adopt the average of the PBMC and the MCMC uncertainties, ζ_i, scaled by 0.887. We stress that these values refer to the point-to-point uncertainties only, which determine the shape of the spectrum, not to the absolute eclipse depths, which can be affected by additional systematic effects, as discussed below.

We examine the possibility for placing an upper limit on the time variability of the planet's emission in different wavelengths. However, the uncertainties of the individual eclipse observations as a function of time are too large to place meaningful constraints on this quantity in a way similar to the 8 μm multi-epoch photometric study by Agol et al. (2010).

Another interesting caveat is that in reality the spectral coverage of the 7–14 μm IRS spectra extends slightly beyond 15 μm. The intensities observed at wavelengths above $\sim$13.5 μm are affected by the well-documented "teardrop" effect protosol,[12] which causes the spectral trail to appear "lumpy" on the images, like a teardrop. The cause of this systematic effect is either light leakage or defects in the detector or optics. There is no reliable correction, and therefore we exclude all eclipse depths at wavelength longer than $\sim$13.5 μm from our results. We summarize the 56 final eclipse depths and their uncertainties in Table 3.

4. DISCUSSION

4.1. Comparison to Previous Studies

The eclipse depths we measure via secondary eclipse spectroscopy appear systematically below the broadband photometry measured by Charbonneau et al. (2008) and Agol et al. (2010). Grillmair et al. (2008) also find deeper eclipses than this study (by $\sim$20%), especially below wavelengths of $\sim$7.5 μm. As seen in the top panel of Figure 7, the difference decreases and becomes marginal for wavelengths above $\sim$7.5 μm.

There are several possible explanations for this discrepancy. For instance, the host star HD 189733 has some variability. Then, it is possible that the observations presented by Charbonneau et al. (2008) and Grillmair et al. (2008) were mostly performed near stellar flux minima, leading to deeper apparent eclipses, while the ones analyzed here in addition could have been observed near stellar flux maxima causing the eclipses to appear shallower. The Agol et al. (2010) point is based on six separate eclipse depth measurements at 8 μm and lies closer to our results than the Charbonneau et al. (2008) result at this wavelength. The variability of HD 189733 is caused by large stellar spots that cover $\sim$1%–2% of the stellar surface (Henry & Winn 2008) and are $\sim$1000 K cooler than the rest of the photosphere (Pont et al. 2007). They lead to a flux variability of around $\sim$1%–2%

[12] http://irsa.ipac.caltech.edu/data/SPITZER/docs/irs/features/#8_SL1_14u_Teardrop.

Table 3
Secondary Eclipse Spectroscopy Results

Wavelength (μm)	Eclipse Depth (%)
IRS 5–7 μm Mode	
5.46	0.211 ± 0.031
5.55	0.195 ± 0.036
5.65	0.195 ± 0.026
5.74	0.222 ± 0.023
5.83	0.251 ± 0.037
5.92	0.190 ± 0.029
6.02	0.198 ± 0.023
6.11	0.214 ± 0.028
6.20	0.256 ± 0.039
6.30	0.243 ± 0.027
6.39	0.227 ± 0.026
6.48	0.252 ± 0.023
6.58	0.192 ± 0.029
6.67	0.194 ± 0.027
6.76	0.194 ± 0.028
6.85	0.239 ± 0.030
6.95	0.189 ± 0.032
7.04	0.240 ± 0.031
7.13	0.226 ± 0.029
7.23	0.292 ± 0.038
7.32	0.207 ± 0.041
7.41	0.318 ± 0.049
7.51	0.227 ± 0.068
IRS 7–14 μm Mode	
7.53	0.314 ± 0.050
7.66	0.287 ± 0.036
7.84	0.257 ± 0.034
8.03	0.296 ± 0.034
8.22	0.266 ± 0.028
8.40	0.324 ± 0.034
8.59	0.317 ± 0.032
8.77	0.343 ± 0.034
8.96	0.318 ± 0.036
9.15	0.366 ± 0.037
9.33	0.314 ± 0.030
9.52	0.333 ± 0.035
9.71	0.382 ± 0.038
9.89	0.362 ± 0.042
10.08	0.377 ± 0.035
10.27	0.420 ± 0.036
10.45	0.340 ± 0.038
10.64	0.331 ± 0.050
10.82	0.418 ± 0.045
11.01	0.394 ± 0.038
11.20	0.417 ± 0.041
11.38	0.401 ± 0.048
11.57	0.357 ± 0.035
11.76	0.408 ± 0.046
11.94	0.416 ± 0.041
12.13	0.367 ± 0.044
12.31	0.522 ± 0.040
12.50	0.473 ± 0.047
12.69	0.443 ± 0.042
12.87	0.383 ± 0.047
13.06	0.378 ± 0.045
13.24	0.461 ± 0.075
13.43	0.365 ± 0.046

in the visible. Assuming that the spots and the photosphere have blackbody fluxes, a stellar variability of 1.5% at 0.5 μm corresponds to variability of about 0.4%–0.6% between 3.6 and 16 μm. This is insufficient to explain the discrepancy between

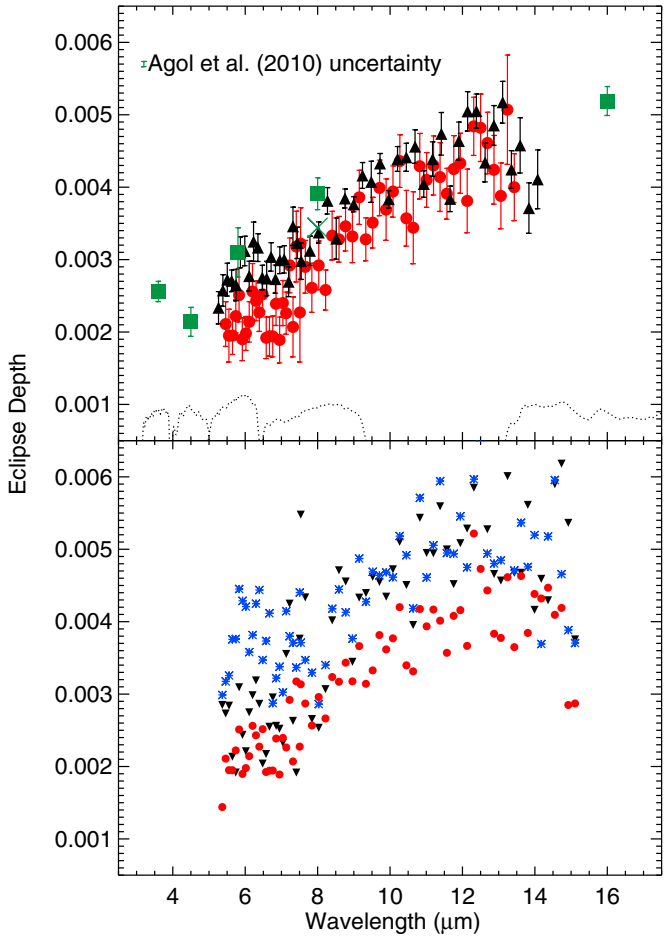

Figure 7. Top panel: comparison between the eclipse depth spectrum derived here (filled red circles) with the results from Grillmair et al. (2008, black triangles) and the Charbonneau et al. (2008) and Agol et al. (2010) secondary eclipse photometry (green squares and green crosses, respectively). The size of the uncertainty in the Agol et al. (2010) measurement is indicated in the upper left corner. Bottom panel: comparison between our adopted results (filled red circles) and analyses of our data using a third-order polynomial ramp correction function (black upside-down triangles; Grillmair et al. 2007) and the Charbonneau et al. (2008) logarithm plus quadratic ramp correction function (blue asterisks), discussed in Section 3.2.1. The results based on the ramp functions in Equations (3), (6), and (7) show similar offsets from the results we adopt, and are not shown here for clarity. The choice of ramp function affects the absolute eclipse depths, but not their overall shapes. Consistently with the previous studies, both of the ramp functions used in the two previous studies yield deeper eclipses than the exponential ramp adopted here.

(A color version of this figure is available in the online journal.)

the results presented here and the Charbonneau et al. (2008) and Grillmair et al. (2008) findings.

Another possibility is that the discrepancy is due to a systematic difference between the sawtooth correction algorithms used in Grillmair et al. (2008) and those used in this study. Any bias or offset in the sawtooth determination can lead to removing or adding to the eclipse depths. This is a more significant problem at the shorter wavelengths, between 5 and $\sim$7.5 μm, where the sawtooth amplitude is higher, perhaps explaining why the difference becomes small at the long wavelengths. However, this effect is likely to be relatively small.

The most likely reason for this is the choice of a ramp correction function (Section 3.2.1). The way the ramp correction function bends at the time of eclipse can be critical for the absolute scale of the eclipse depth, as seen in the bottom panel of Figure 7. While the shape of the spectrum is unchanged regardless of the choice of ramp correction function, substituting

the physically motivated ramp correction function, we adopt in our analysis with the functions used by Grillmair et al. (2007) and Charbonneau et al. (2008) a third-order polynomial (Equation (5)) and a logarithm plus quadratic (Equation (4)), respectively, yields results comparable with the Grillmair et al. (2008) absolute depths. We suggest that a consistent re-analysis of the *Spitzer*/IRAC photometry utilizing the single exponent ramp correction function could yield somewhat lower eclipse depths and could lead to a better agreement between the photometric points and the updated spectroscopy presented here. However, this analysis lies outside the scope of this work.

The offset in absolute eclipse depth, likely caused by the choice of ramp correction, is not critical for our results because it appears to change slowly with wavelength, thus generally preserving the shape of the spectrum. It does, however, limit our ability to draw conclusions about the dayside-to-nightside energy transfer efficiency, since this is the property that determines the overall dayside planetary flux levels. However, an examination of the shape of the pressure–temperature (*P–T*) atmospheric profile is still possible.

4.2. Comparison to Emergent Spectrum Models

In order to study the atmospheric properties of HD 189733b, we compare the measured emergent spectrum to models. We utilize a simple and computationally efficient radiative transfer code developed by Richardson et al. (2003), which is sufficient to retrieve the basic thermal properties of the atmosphere. The atmospheric composition is assumed to be solar, but the only elements that are explicitly tracked are H, He, C, and O. Line opacities of CO, CH_4, and H_2O are included, as well as collision-induced absorption of H_2–H_2 (e.g., Borysow & Frommhold 1990; Borysow 2002) and H_2–He (Jørgensen et al. 2000). The molecular mixing ratios are computed based on the approximate method given in the Appendix of Burrows & Sharp (1999). The code adopts the water line wavelengths and strengths from Partridge & Schwenke (1997), the CO lines are from Goorvitch (1994), and the CH_4 lines are from the High-resolution Transmission Molecular Absorption Database (HITRAN; Rothman et al. 1998). The line opacities are computed explicitly using Voigt profile line shapes with pressure broadening coefficient of $0.1\,\mathrm{cm^{-1}\,atm^{-1}}$. If the water lines are considered individually, since there are hundreds of millions of them and many overlap at a given wavelength, the model calculations become computationally expensive. To mitigate this issue, we bin the water lines, to reduce the number of them that need to be included at a given wavelength. The individual water line strengths are dependent on temperature, and so is the strength of the binned lines. Therefore, we compute the binned line strengths for six different temperatures between 300 and 3000 K, and we approximate the log of the total strength of the binned lines as a quadratic function of the logarithm of the temperature. Since the quadratic is used in log-log space, it is very well able to capture the variation of the total binned line strength with temperature. To confirm this, we have inspected visually a number of these fits. The CO and CH_4 lines are less numerous and do not require binning.

The Richardson et al. (2003) code does not incorporate hazes or clouds. Hazes have been recently observed in the UV dayside spectrum of HD 189733b (Evans et al. 2013) but appear to lose importance toward visible wavelengths. Hazes also appear to play an important role in the formation of the transmission spectrum of HD 189733b (which is measured near the terminator, not on the dayside, like the emission spectrum) at wavelengths below $\sim$2.5 μm (e.g., Pont et al. 2008; Gibson et al. 2012). Gibson et al. (2012) suggest the existence of a Rayleigh-scattering haze in the planet's atmosphere, but this would be expected to become transparent at longer wavelengths. In addition, the light paths through the atmosphere for the emission spectrum are near vertical, much shorter than the planet-grazing light paths observed during transit spectroscopy. Hence, small amounts of atmospheric haze will have a much smaller cumulative effect on the dayside emission spectrum than in the transmission spectrum of the planet. Thus, it is unlikely that hazes are essential for modeling the infrared dayside emission spectrum of the planet.

In order to calculate the model star-planet contrast, or secondary eclipse depth, we adopt a Kurucz model for the host star (Kurucz 1979) and use its predicted brightness at a given wavelength to divide the model planet brightness. This yields the model eclipse depth at this wavelength. The most important input for the Richardson et al. (2003) model is the *P–T*, and we calculate the model emission spectrum for several *P–T* profiles.

As Burrows (2014) argues, many of the exoplanet characterization results to date are fragile and possibly misleading. Thus, we compare our results on the well-studied HD 189733b to models with isothermal, gray, and full non-gray hydrostatic equilibrium atmospheres in order to test what we know truly robustly. Eliminating the simplest models reliably would ensure that we do not infer more atmospheric information based on a model than what is justified by the data.

4.3. Isothermal Model Atmosphere

The simplest possible model atmospheric *P–T* profile is isothermal, which produces a blackbody emission spectrum. If the measured eclipse depths for this planet are consistent with blackbody, this would be a strong indication that we can extract very little information about this planet's atmosphere from our data. A comparison between a blackbody planet and our derived spectrum is presented in Figure 8.

While for wavelengths of $\lambda \gtrsim 8\,\mu$m the isothermal atmosphere prediction matches the observations, it cannot explain well the "bump"-like feature near 6.3 μm, which is likely caused by water *absorption*. We examine this feature in detail in Section 4.6. The shape of the Charbonneau et al. (2008) photometric spectral energy distribution (SED) is also inconsistent with a blackbody planet, especially at 3.6 μm. However, this photometric study did not have the benefit of modern analysis techniques. In addition, more recent measurements by Knutson et al. (2012) find shallower eclipse depths at 3.6 and 4.5 μm, inconsistent with the corresponding Charbonneau et al. (2008) measurements. Therefore, photometry alone cannot be solely used to rule out the isothermal model.

4.4. Gray Pressure–Temperature Atmospheric Profile

A gray atmosphere, where the gas opacity is equal at all wavelengths, has a more complex *P–T* profile than an isothermal atmosphere (for a detailed discussion of gray atmospheres; see, e.g., Rutten 2003). We adopt the Rosseland mean opacities for solar metallicity given by Freedman et al. (2008) for ultracool brown dwarfs and exoplanets for temperatures between 75 and 4000 K, and interpolate between the values provided in their tables as necessary. We begin with an "initial guess" *P–T* profile (isothermal one) and use it to estimate the run of Rosseland opacity with pressure. We then calculate the optical depth and, hence, temperature as a function of pressure. Using this, we re-estimate the Rosseland opacity as a function of pressure, and

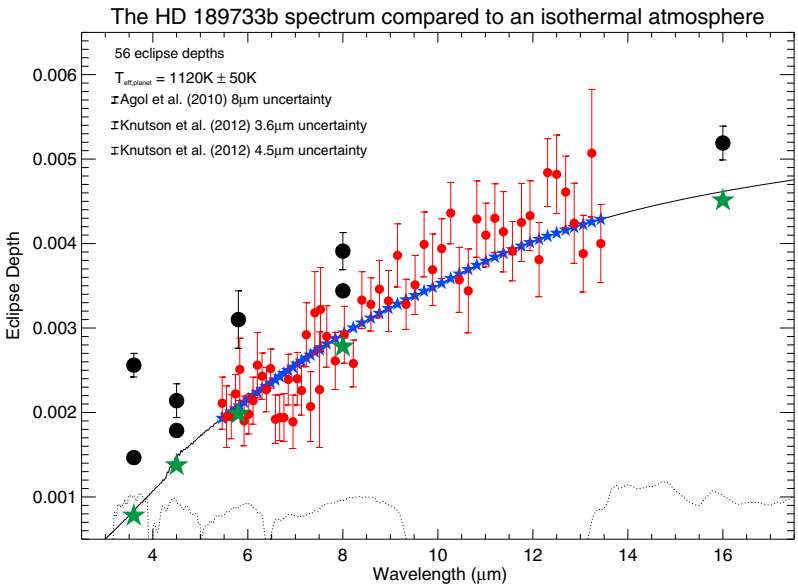

Figure 8. We compare the observed planet-to-star contrast of HD 189733 (large red points) with 56 wavelengths with a prediction based on a blackbody planet, i.e., isothermal atmosphere (solid black line). The broadband secondary eclipse depths measured by Charbonneau et al. (2008; black circles with error bars), Knutson et al. (2012), and Agol et al. (2010; black circles with no error bars at 3.6, 4.5, and 8 μm, respectively) using *Spitzer* IRAC and MIPS are also plotted. The Agol et al. (2010) measurement at 8 μm represents an average of six separate eclipse depth measurements. The size of the uncertainties of the Agol et al. (2010) and Knutson et al. (2012) results is indicated in the upper left for clarity. We show the band integrated contrasts predicted for a blackbody planet for the IRS data (blue stars) and the IRAC and MIPS bandpasses (green stars). The black dotted lines indicate the transmission functions of *Spitzer* IRAC and MIPS. We indicate the effective temperature of the planet assuming an isothermal atmosphere in the upper left, with uncertainties derived from the χ^2 distribution and based only on the spectroscopic data points. For this fit, the minimum $\chi^2_{\rm red} = 1.02$. This value does not include the photometric points. The only degree of freedom here is the planetary temperature.

(A color version of this figure is available in the online journal.)

repeat the process until the run of temperatures with pressure converges. The effective temperature of the planet, $T_{\rm eff}$, is given by the temperature at optical depth, $\tau = 2/3$, and this is a model input parameter, represented by the temperature of the initial isothermal profile guess. Its value sets absolute values of the *P*–*T* profile, as opposed to the profile's shape, which is set by the gray atmosphere. We experiment with plugging the gray *P*–*T* profile for a range of values for $T_{\rm eff}$ in the Richardson et al. (2003) radiative transfer code, where the atmosphere is *not* assumed to be gray, and examine the resulting planetary emission spectra. Using χ^2 minimization, we find that $T_{\rm eff} = 1300\,{\rm K} \pm 100\,{\rm K}$ yields an emergent spectrum closest to the observed one. The uncertainties on this quantity are based on the χ^2 distribution of the spectroscopic data, assuming that only $T_{\rm eff}$ is a free parameter. Here we exclude the photometric data points, since it is difficult to estimate their systematic offset and thus to assign them appropriate weights. We show the results from this calculation in Figures 9 and 10. As with the isothermal model, a gray atmosphere *P*–*T* profile appears to be unlikely based on the shape of the Charbonneau et al. (2008) broadband SED, especially the 3.6 μm eclipse depth, despite lack of updated reduction of these data. However, a gray atmosphere appears to be consistent with the more recent Knutson et al. (2012) photometric data points at 3.6 and 4.5 μm. As with the isothermal model, the gray atmosphere fails to account for the secondary eclipse spectrum near the 6.3 μm "bump," discussed in Section 4.6.

4.5. The Burrows Atmosphere Model

In their study, Grillmair et al. (2008) compare their results to a model developed by Burrows et al. (2007, 2008) that adopts the chemical equilibrium and opacities information computed by Burrows & Sharp (1999) and Sharp & Burrows (2007).

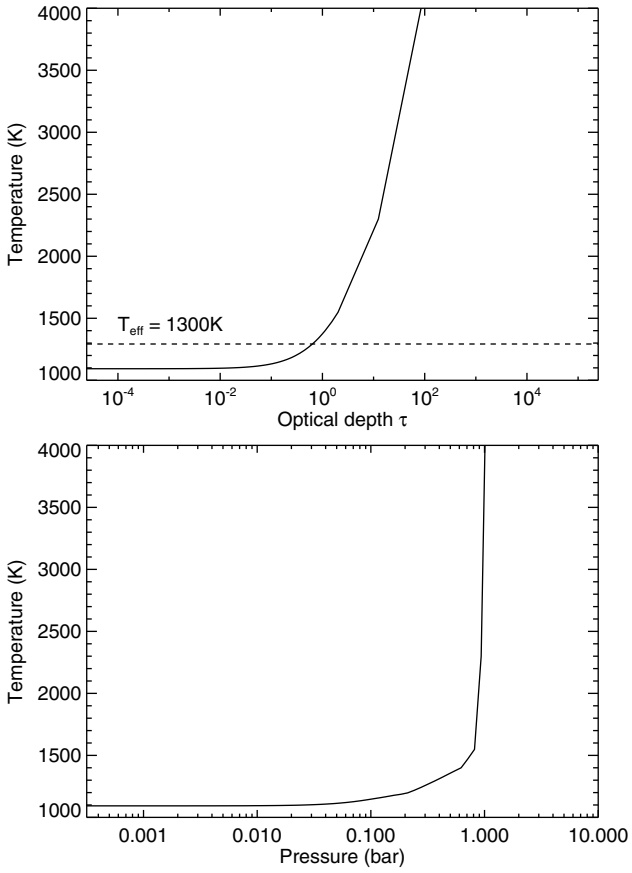

Figure 9. Calculated temperature as a function of optical depth (upper panel) and pressure (lower panel), assuming a gray atmosphere with adopted Rosseland mean opacities from Freedman et al. (2008). The dashed line indicates the effective temperature of the planet, $T_{\rm eff}$, at $\tau = 2/3$.

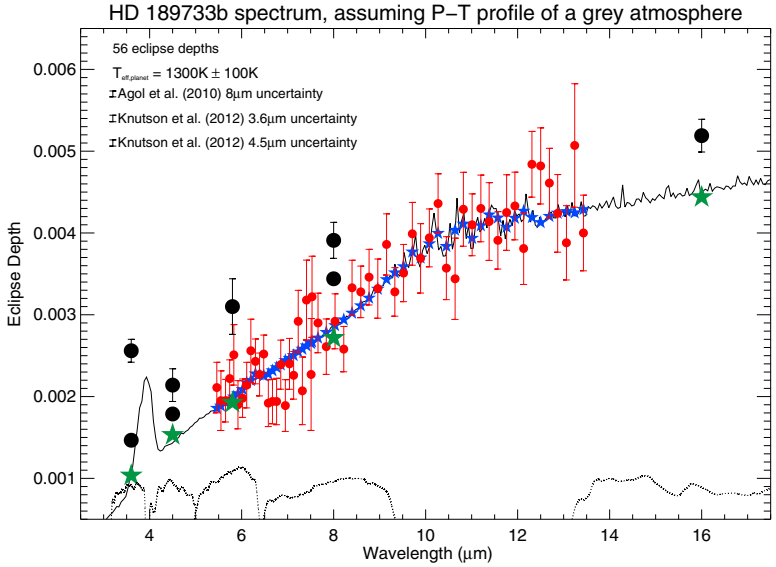

Figure 10. Similar to Figure 8, but here we show a Richardson et al. (2003) model based on the gray P–T profile computed in Section 4.4 (solid black line) compared to the observed results—large red points with error bars (spectroscopy) and black circles (photometry). The band integrated contrasts predicted for a gray atmosphere are represented by blue stars (spectroscopy) and green stars (photometry). As in Figure 8, the effective temperature of the planet, $T_{\rm eff}$, is shown in the upper left, with uncertainties based on the χ^2 distribution, taking into account only the spectroscopic data. A gray atmosphere P–T profile is ruled out because it fails to account for the 6.3 μm bump and for the Charbonneau et al. (2008) 3.6 μm eclipse depth. The minimum reduced χ^2 value for this fit, assuming that only $T_{\rm eff}$ is free, is $\chi^2_{\rm red} = 0.77$. Again, we only used the spectroscopic points for the fit and the $\chi^2_{\rm red}$ calculation. As in the blackbody case, here there is only one degree of freedom, $T_{\rm eff}$.
(A color version of this figure is available in the online journal.)

The model relies on fully non-gray radiative opacities with layer-by-layer radiative equilibrium and chemical equilibrium using an extensive set of molecular and atomic abundances. The model also includes a parameterized generic stratospheric flux absorber that can cause a temperature inversion and a parameterized dayside-to-nightside heat transfer efficiency. We compare the Burrows model adopted by Grillmair et al. (2008) to our measurements. We also use the Burrows-derived P–T profile as input to the Richardson et al. (2003) radiative transfer code and compare the result with our observations. The Burrows model presented here is similar to the one shown in Grillmair et al. (2008) and has a heat redistribution parameter $P_n = 0.1$ implying that only 10% of the heat absorbed on the planet's dayside is transferred to the nightside (the maximum is $P_n = 0.5$, or 50%). The absorption coefficient of the hypothetical unknown high-altitude absorber is set to $\kappa_{\rm abs} = 0.020\,{\rm cm}^2\,{\rm g}^{-1}$, meaning that the planet is assumed to have no or negligible stratosphere and relatively inefficient heat transfer to its nightside. Later studies have also found no evidence for inversion (Madhusudhan & Seager 2009; Swain et al. 2009b), and the planet is typically considered to have a non-inverted atmosphere (e.g., Knutson et al. 2010). The Burrows model cannot be ruled out based on our analysis. We present the P–T pressure for this model in Figure 11 and a comparison between the emission from the Burrows atmosphere to our observed eclipse depths in Figure 12.

4.6. Discussion of Results and Implications for the Atmosphere

Grillmair et al. (2008) describe a "bump" in their spectrum near 6.2 μm that they attribute to the opacity minimum between the P and R branches of the ν_2 band at 6.27 μm (the fundamental vibrational bending mode of water). These authors also notice a very tenuous rise in emission at 5.9 μm that they are unable to identify. These features are preserved after the inclusion of the previously never analyzed observations in this study (Figure 13). Even though the flux near 5.9 μm still appears higher than that

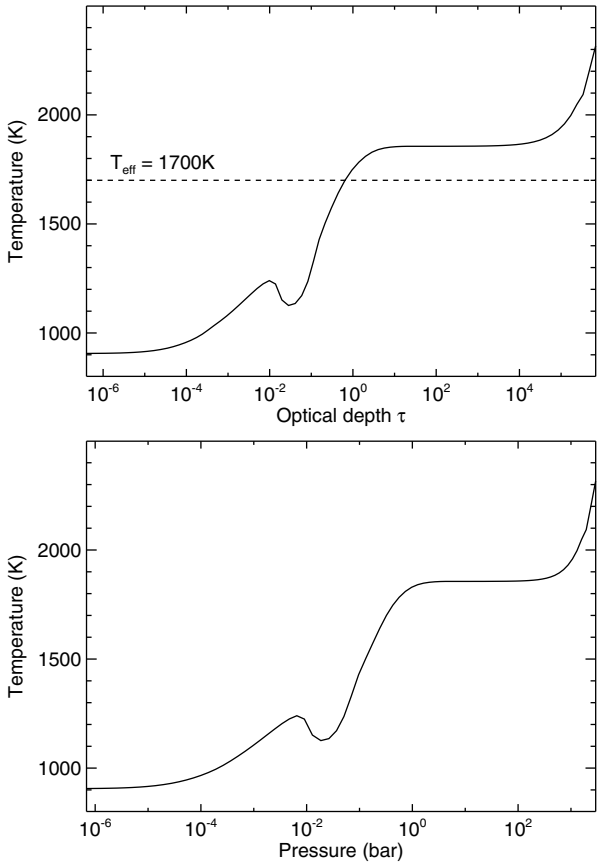

Figure 11. Temperature as a function of optical depth based on the Rosseland opacity (upper panel) and pressure (lower panel) used to calculate the Burrows model similar to the one adopted by Grillmair et al. (2008). As in Figure 9, the dashed line indicates the effective temperature of the planet, $T_{\rm eff}$. Here the high-altitude absorber is assumed to be relatively unimportant for the emission spectrum, with $\kappa_{\rm abs} = 0.02\,{\rm cm}^2\,{\rm g}^{-1}$, but it still causes a small temperature inversion near pressure of 0.01 bar.

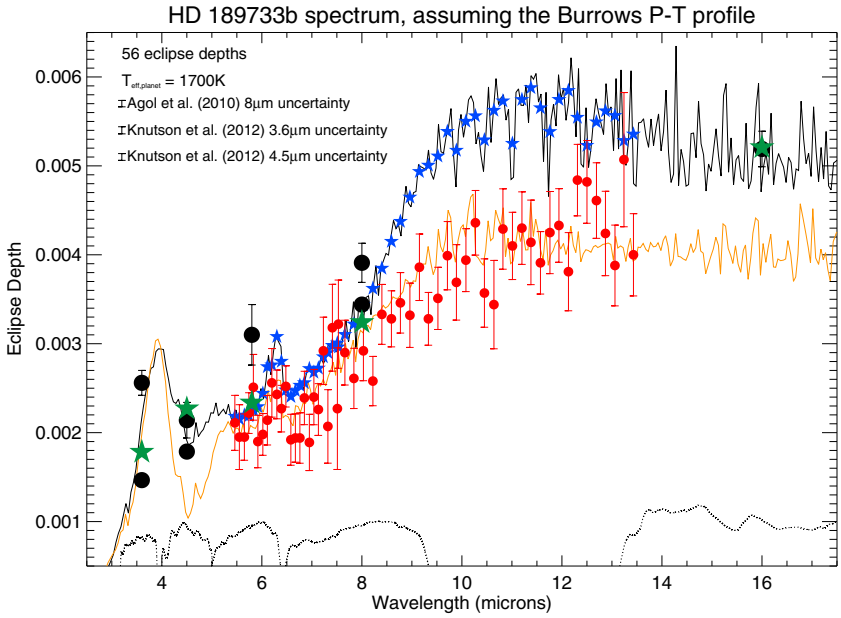

Figure 12. Similar to Figures 8 and 10, but here we compare our observed planet-star contrast (large red points) to two models produced based on the *P–T* profile shown in Figure 11. These models are prohibitively computationally intensive to calculate, making statistical fitting impractical. Therefore, we do not fit them to the data, but we plot them with the observed data for comparative purposes. The planetary emission model similar to the one adopted by Grillmair et al. (2008) is calculated based on the studies by Burrows et al. (2007, 2008; solid orange line). The black line is a simpler Richardson et al. (2003) model based on the Grillmair et al. (2008) *P–T* profile that we calculate. As before, the star symbols represent a band-integrated version of this model. We do not claim that either of these two models is preferred, but they give a better match to the shape of the planetary spectrum, compared to the isothermal and gray models, even though they both predict deeper eclipses than observed. For the better matching Burrows et al. (2007, 2008) model, we calculate the non-optimized $\chi^2 = 406.8$, based on the 56 IRS eclipse depths (this value does not take the *Spitzer* IRAC photometry into account). While there are numerous input parameters for the Burrows models, they are all kept fixed here. Thus, we do not calculate the reduced χ^2 value.

(A color version of this figure is available in the online journal.)

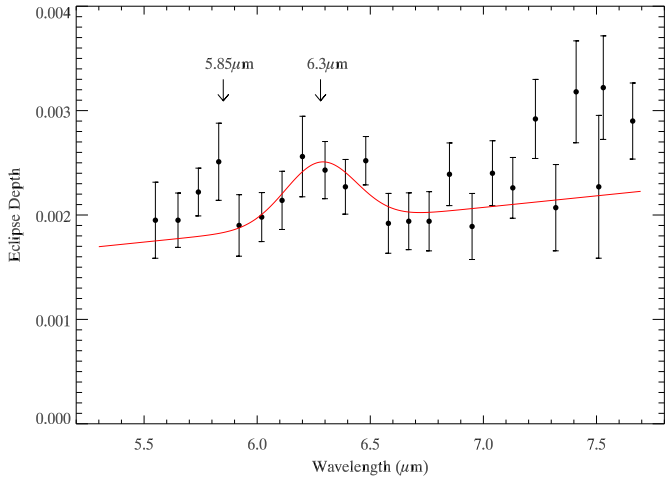

Figure 13. Section of the observed secondary eclipse spectrum of HD 189733b near the 6 μm water feature. This apparent "bump" in the spectrum is caused by an opacity minimum near 6.27 μm between two *absorption* features of water on either side (see text for details). We have overplotted a Gaussian curve centered at 6.32 μm with width of 0.15 μm. We also mark the tenuous unidentified emission bump near 5.9 μm (Grillmair et al. 2008), which we consider to be most likely a noise artifact.

(A color version of this figure is available in the online journal.)

of its neighbors, the uncertainties are too large to be able to claim even a tentative detection.

In order to test the robustness of the detection of the 6 μm water feature, we focus on the 5.9–7.0 μm range of the results. We fit a Gaussian function added to a sloped line and measure an amplitude of the feature of 0.00058 in units of contrast and a wavelength of the maximum at 6.28 μm, with a Gaussian width, $\sigma = 0.15\,\mu$m. We run a Monte Carlo simulation of 10,000 spectra in this range, by drawing random numbers from Gaussian distributions with means equal to the fitted straight line but with the Gaussian feature removed and with widths equal to the observed eclipse depth uncertainties. We fit every simulated spectrum with a Gaussian in the same way that we fit the observed water feature and find that only ~3% of the simulated data sets exhibit peaks similar to that of the real spectrum—with amplitudes greater than 0.0001 and widths between 0.1 and 0.2 μm.

Therefore, we rule out a smooth spectrum for this wavelength range and thus reject the isothermal model for the atmosphere at the 97% level, based solely on the IRS data. Our analysis is in agreement with the conclusions of (Grillmair et al. 2008), who detect the 6 μm water feature at the 95% level.

The shape of the gray atmosphere model is similar to the shape of the blackbody spectrum at 5–7 μm wavelength range, especially at low wavelength resolution, as in our data (Figure 10). Therefore, the gray atmosphere model can be rejected at the 97% level with the same argument as the isothermal model—it is in disagreement with the detection of the 6 μm feature.

The Burrows and Richardson et al. (2003) models based on the *P–T* profile similar to the one adopted in Grillmair et al. (2008) provide the best match to the shape of the observed spectrum, especially near the 6 μm water feature. However, both of these predict deeper eclipses than the ones we observe. This makes the Burrows-based models more difficult to reject outright, even though they do not provide a perfect match for the measurements.

For completeness, we compare the χ^2 statistics produced by the three models (excluding the broadband eclipses, whose

weights in the calculation would be nontrivial). The isothermal atmosphere model produces $\chi^2 = 54.9$, larger than the value of $\chi^2 = 41.4$ produced by the gray atmosphere model (our spectrum consists of 56 points). For the full Burrows radiative transfer model, $\chi^2 = 406.8$. While the χ^2 values strongly favor the simpler models, they are a poor match to the observed data near 6 μm, and we have not minimized the χ^2 parameter for the Burrows model as this is impractical. More importantly, the χ^2 statistic is very sensitive to the absolute eclipse depth values compared to the model, so it is optimal for testing the absolute eclipse depths, not the shape of the spectrum, on which we have focused here. Since a true fit is not feasible, we attempt applying absolute shifts to the data and comparing it to the Burrows model, until the χ^2 is minimized. We find that $\chi^2 = 160.2$. However, drawing any conclusions out of this (e.g., that the model is rejected) is difficult, because a true model fit with realistic free parameters is needed in order to show this definitely.

While the combination of photometry and spectroscopy proves the isothermal or gray atmosphere explanation for the observations unlikely, the rejection hinges on data within a small wavelength range. This underscores the necessity of large wavelength coverage and high spectral resolution required to make definitive measurements of the properties of exoplanet atmospheres, as suggested by the model retrieval efforts by Lee et al. (2012) and Line et al. (2012, 2013). While the Burrows model cannot be rejected outright, it is possible that additional data in other wavelengths may show that it is an insufficient explanation of the atmosphere of this hot Jupiter.

5. CONCLUSION

Our analysis of the *Spitzer*/IRS secondary eclipse time series spectroscopy of HD 189733b has largely confirmed the results from the original Grillmair et al. (2008) study, despite the fact that it did not include many of the observations available to us. While we confirm the detection of the 6 μm water feature, we find that broad wavelength coverage and high spectral resolution are essential for the studies of exoplanet atmospheres, since we are only able to reject the simplistic isothermal and gray atmosphere models based on only several out of a total of ~60 eclipse depth measurements. The fact that our results are in broad agreement with Grillmair et al. (2008) is an important confirmation that many of the systematics correction techniques the exoplanet community has been employing for *Spitzer* observations are robust and can be applied to the next generation of space-based exoplanet studies as necessary. The offset between the Charbonneau et al. (2008) photometry and the updated spectroscopy in this work, as well as the photometry in Knutson et al. (2012), should be examined further, in particular with a consistent updated analysis of all secondary eclipse photometric data sets available for this planet. The offset is likely related to the choice of ramp correction function. HD 189733b is the exoplanet with the most thoroughly studied atmosphere to date, and our results enhance this achievement, despite the fact that plenty of questions about its atmosphere's structure and composition remain open. These questions could be answered by future studies in the context of new and archival observations.

We thank the Spitzer Science Center HelpDesk and Sean Carey for their valuable assistance in exploring alternatives for the correction of the systematic effects caused by the telescope pointing jitter. We thank David Charbonneau and Jonathan Fortney for valuable discussions. We thank the referee for the detailed review and many thoughtful suggestions. This work is based on observations made with the *Spitzer Space Telescope*, obtained from the NASA/IPAC Infrared Science Archive, both of which are operated by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Support for this work was provided by NASA through an award issued by JPL/Caltech. This research has made use of NASA's Astrophysics Data System. This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org.

REFERENCES

Agol, E., Cowan, N. B., Knutson, H. A., et al. 2010, ApJ, 721, 1861
Barstow, J. K., Aigrain, S., Irwin, P. G. J., Fletcher, L. N., & Lee, J.-M. 2013, MNRAS, 434, 2616
Beerer, I. M., Knutson, H. A., Burrows, A., et al. 2011, ApJ, 727, 23
Birkby, J. L., Ce Kok, R. J., Brogi, M., et al. 2013, MNRAS, 436, L35
Borysow, A. 2002, A&A, 390, 779
Borysow, A., & Frommhold, L. 1990, ApJL, 348, L41
Burrows, A. 2014, PNAS, 111, 12601
Burrows, A., Budaj, J., & Hubeny, I. 2008, ApJ, 678, 1436
Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, ApJL, 668, L171
Burrows, A., & Sharp, C. M. 1999, ApJ, 512, 843
Charbonneau, D., Allen, L. E., Megeath, S. T., et al. 2005, ApJ, 626, 523
Charbonneau, D., Knutson, H. A., Barman, T., et al. 2008, ApJ, 686, 1341
Correia, A. C. M., & Laskar, J. 2010, in Exoplanets, ed. S. Seager (Tucson, AZ: Univ. Arizona Press), 239
Cowan, N. B., & Agol, E. 2011, ApJ, 729, 54
Crossfield, I. J. M., Hansen, B. M. S., & Barman, T. 2012, ApJ, 746, 46
Deming, D., Harrington, J., Seager, S., & Richardson, L. J. 2006, ApJ, 644, 560
Deming, D., Knutson, H., Agol, E., et al. 2011, ApJ, 726, 95
Deming, D., Seager, S., Richardson, L. J., & Harrington, J. 2005, Natur, 434, 740
Désert, J.-M., Charbonneau, D., Fortney, J. J., et al. 2011, ApJS, 197, 11
Eastman, J., Siverd, R., & Gaudi, B. S. 2010, PASP, 122, 935
Evans, T. M., Pont, F., Sing, D. K., et al. 2013, ApJL, 772, L16
Ford, E. B. 2005, AJ, 129, 1706
Ford, E. B. 2006, ApJ, 642, 505
Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419
Fortney, J. J., Saumon, D., Marley, M. S., Lodders, K., & Freedman, R. S. 2006, ApJ, 642, 495
Freedman, R. S., Marley, M. S., & Lodders, K. 2008, ApJS, 174, 504
Gibson, N. P., Aigrain, S., Pont, F., et al. 2012, MNRAS, 422, 753
Gillon, M., Demory, B.-O., Barman, T., et al. 2007, A&A, 471, L51
Goorvitch, D. 1994, ApJS, 95, 535
Grillmair, C. J., Burrows, A., Charbonneau, D., et al. 2008, Natur, 456, 767
Grillmair, C. J., Charbonneau, D., et al. 2007, ApJL, 658, L115
Henry, G. W., & Winn, J. N. 2008, AJ, 135, 68
Horne, K. 1986, PASP, 98, 609
Houck, J. R., Roellig, T. L., van Cleve, J., et al. 2004, ApJS, 154, 18
Hubeny, I., Burrows, A., & Sudarsky, D. 2003, ApJ, 594, 1011
Jackson, B., Greenberg, R., & Barnes, R. 2008, ApJ, 678, 1396
Jørgensen, U. G., Hammer, D., Borysow, A., & Falkesgaard, J. 2000, A&A, 361, 283
Knutson, H. A, Charbonneau, D., Allen, L. E., et al. 2007, Natur, 447, 183
Knutson, H. A., Charbonneau, D., Allen, L. E., Burrows, A., & Megeath, S. T. 2008, ApJ, 673, 526
Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, ApJ, 720, 1569
Knutson, H. A., Lewis, N., Fortney, J. J., et al. 2012, ApJ, 754, 22
Kurucz, R. L. 1979, ApJS, 40, 1
Lee, J.-M., Fletcher, L. N., & Irwin, P. G. J. 2012, MNRAS, 420, 170
Lewis, N. K., Knutson, H. A., Showman, A. P., et al. 2013, ApJ, 766, 95
Line, M. R., Wolf, A. S., Zhang, X., et al. 2013, ApJ, 775, 137
Line, M. R., Zhang, X., Vasisht, G., et al. 2012, ApJ, 749, 93
Machalek, P., McCullough, P. R., Burrows, A., et al. 2009, ApJ, 701, 514
Madhusudhan, N., & Seager, S. 2009, ApJ, 707, 24
Mandel, K., & Agol, E. 2002, ApJL, 580, L171
Markwardt, C. B. 2009, in ASP Conf. Ser. 411, Astronomical Data Analysis Software and Systems XVIII, ed. D. A. Bohlender, D. Durand, & P. Dowler (San Francisco, CA: ASP), 251
Parmentier, V., Showman, A. P., & Lian, Y. 2013, A&A, 558, A91

Partridge, H., & Schwenke, D. W. 1997, JChPh, 106, 4618
Pont, F., Gilliland, R. L., Moutou, C., et al. 2007, A&A, 476, 1347
Pont, F., Knutson, H., Gilliland, R. L., Moutou, C., & Charbonneau, D. 2008, MNRAS, 385, 109
Richardson, L. J., Deming, D., Horning, K., Seager, S., & Harrington, J. 2007, Natur, 445, 892
Richardson, L. J., Deming, D., & Seager, S. 2003, ApJ, 597, 581
Rothman, L. S., Rinsland, C. P., Goldman, A., et al. 1998, JQSRT, 60, 665
Rutten, R. J. 2003, Lecture Notes, Radiative Transfer in Stellar Atmospheres (Utrecht: Utrecht Univ.)
Schwarz, G. 1978, AnSta, 6, 461
Sharp, C. M., & Burrows, A. 2007, ApJS, 168, 140
Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, Natur, 464, 1161
Swain, M. R., Bouwman, J., Akeson, R. L., Lawler, S., & Beichman, C. A. 2008a, ApJ, 674, 482
Swain, M. R., Tinetti, G., Vasisht, G., et al. 2009a, ApJ, 704, 1616
Swain, M. R., Vasisht, G., & Tinetti, G. 2008b, Natur, 452, 329
Swain, M. R., Vasisht, G., Tinetti, G., et al. 2009b, ApJL, 690, L114
Tinetti, G., Vidal-Madjar, A., Liang, M.-C., et al. 2007, Natur, 448, 169
Todorov, K. O., Deming, D., Harrington, J., et al. 2010, ApJ, 708, 498
Todorov, K. O., Deming, D., Knutson, H. A., et al. 2012, ApJ, 746, 111
Todorov, K. O., Deming, D., Knutson, H. A., et al. 2013, ApJ, 770, 102
Triaud, A. H. M. J., Queloz, D., Bouchy, F., et al. 2009, A&A, 506, 377
Zahnle, K., Marley, M. S., Freedman, R. S., Lodders, K., & Fortney, J. J. 2009, ApJL, 701, L20